The Impact of Multimodal Large Language Models on Health Care’s Future

Year: 2023 | Citations: 137 | Authors: B. Meskó

Abstract

When large language models (LLMs) were introduced to the public at large in late 2022 with ChatGPT (OpenAI), the interest was unprecedented, with more than 1 billion unique users within 90 days. Until the introduction of Generative Pre-trained Transformer 4 (GPT-4) in March 2023, these LLMs only contained a single mode—text. As medicine is a multimodal discipline, the potential future versions of LLMs that can handle multimodality—meaning that they could interpret and generate not only text but also images, videos, sound, and even comprehensive documents—can be conceptualized as a significant evolution in the field of artificial intelligence (AI). This paper zooms in on the new potential of generative AI, a new form of AI that also includes tools such as LLMs, through the achievement of multimodal inputs of text, images, and speech on health care’s future. We present several futuristic scenarios to illustrate the potential path forward as multimodal LLMs (M-LLMs) could represent the gateway between health care professionals and using AI for medical purposes. It is important to point out, though, that despite the unprecedented potential of generative AI in the form of M-LLMs, the human touch in medicine remains irreplaceable. AI should be seen as a tool that can augment health care professionals rather than replace them. It is also important to consider the human aspects of health care—empathy, understanding, and the doctor-patient relationship—when deploying AI.